

The Challenge of Replication in the Social Sciences

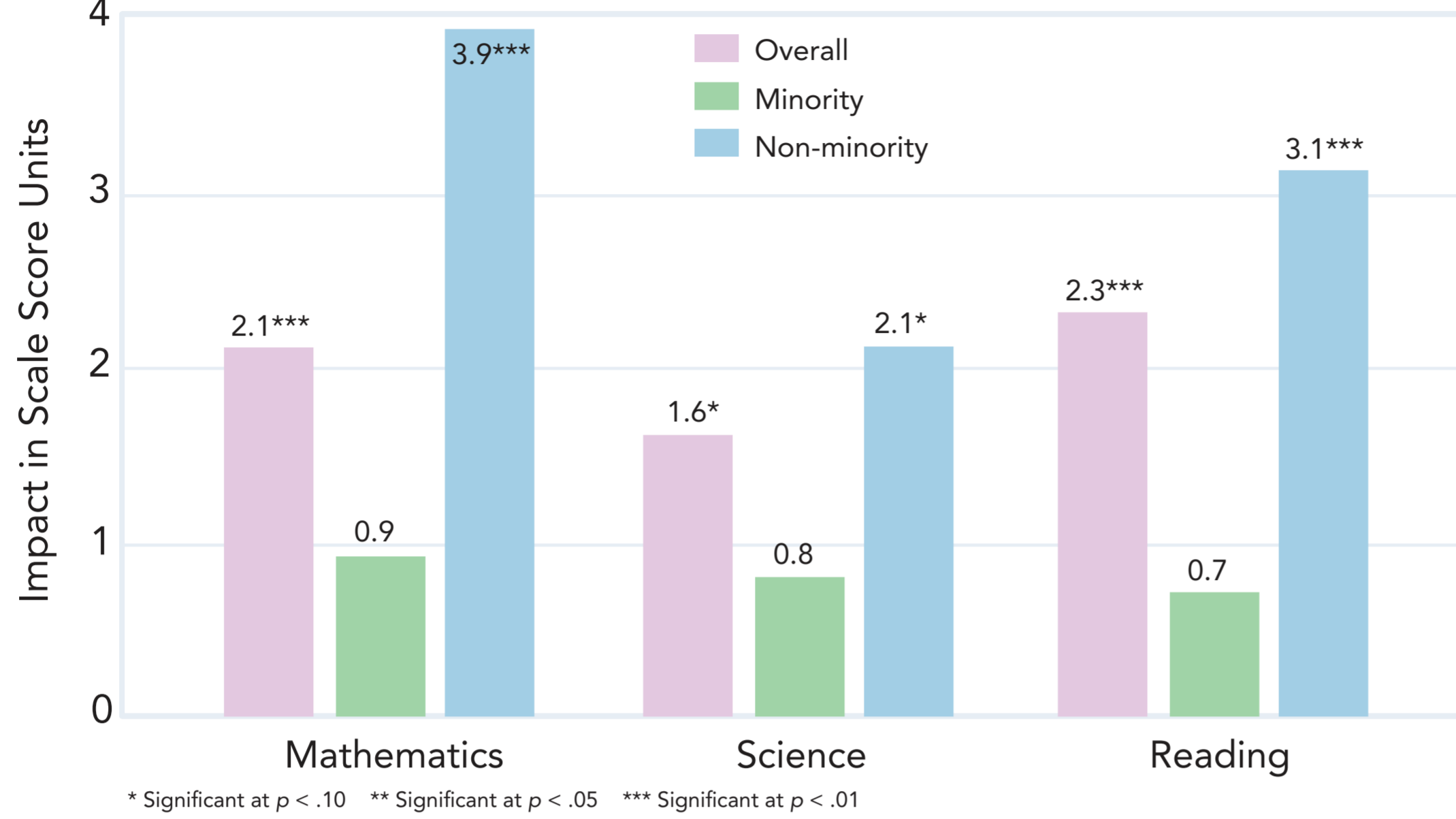
(Taking Randomized Control Trials Further: Prioritizing the Need for Timely and Relevant Results)

Andrew Jaciw, Denis Newman, Boya Ma, Jenna Zacamy, Margit Zsolnay

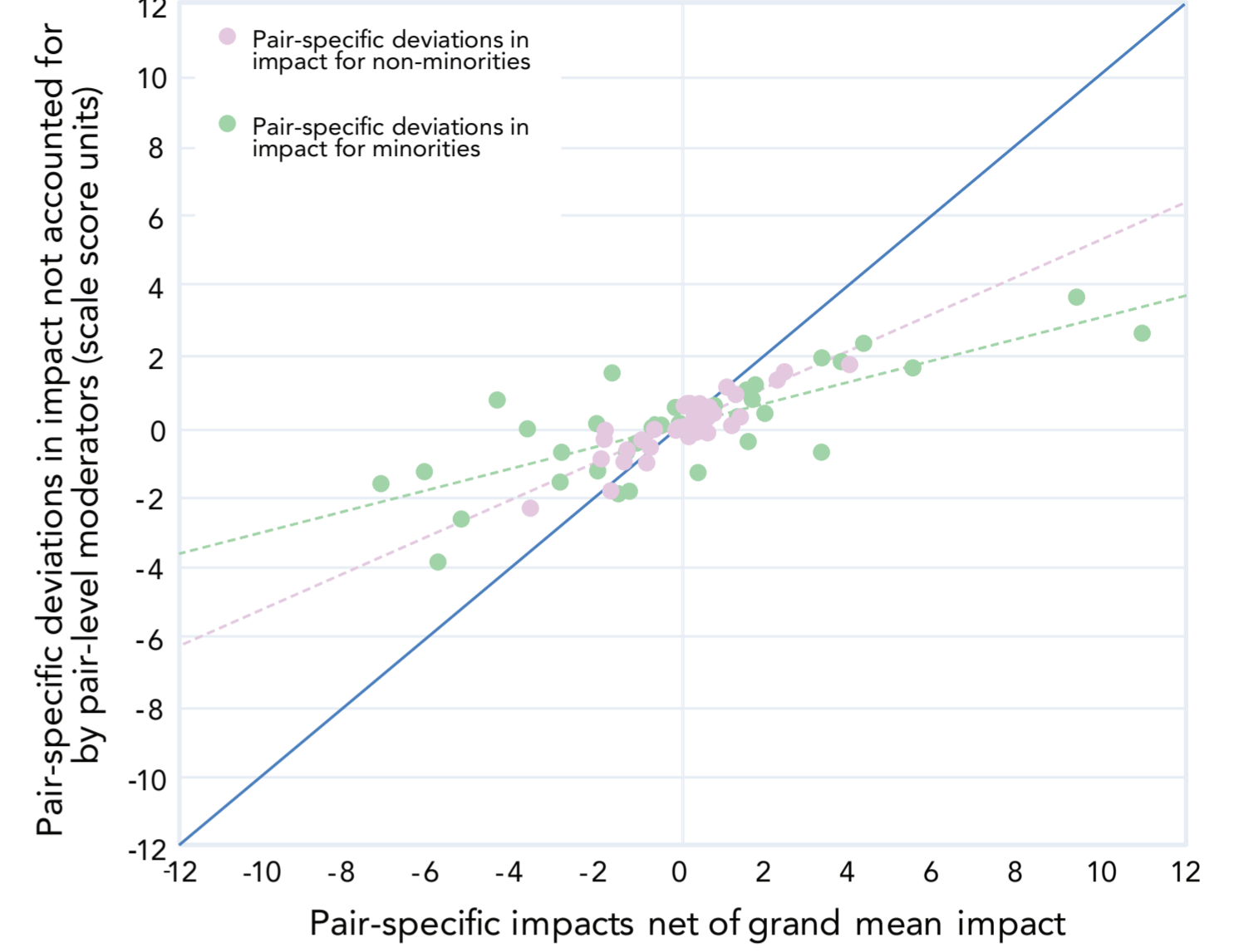


Empirical Education Inc.

IMPACT OF AMSTI OVERALL AND BY MINORITY STATUS FOR THREE SUBJECT AREAS



ESTIMATES OF IMPACTS OF AMSTI BY MATCHED PAIR FOR MINORITIES AND NON-MINORITIES



RISKS AND COSTS OF DECISIONS CONCERNING EXPLORATORY RESULTS

	Exploratory result is true	Exploratory result is false
Assume exploratory result is true (and act to strengthen the program for subgroup)	You avert the cost of depriving subgroup of learning opportunities.	You incur the cost of adapting the implementation of the program unnecessarily.
Assume exploratory result is false (and not act to strengthen the program for subgroup)	You incur the cost of depriving subgroup of learning opportunities.	You avert the cost of adapting the implementation of the program unnecessarily.

- Traditional approaches to research separate analyses and results into primary (confirmatory) and secondary (exploratory) categories.
- Strict standards are applied to confirmatory analyses to limit the chance of drawing false positive conclusions; exploratory analyses give preliminary results that must be confirmed through replication.
- But what happens when exploratory analyses show results that, if real, are consequential? In the AMSTI example, exploratory analyses showed the positive average impact did not extend to Minorities.
- Researchers may be willing to wait (sometimes years) to corroborate the result through another trial; while decision-makers may have to consider the consequences of exploratory results, if they are real.
- The researcher and decision-maker have a different risk calculus
- What is the decision-maker to do when faced with preliminary exploratory, but potentially consequential, results? Ask more of the result if supporting data allow it. With AMSTI we ask: Does the effect persist across types of schools? Does it continue even when we account for the interactions of treatment with other covariates that are not distributed equally across categories of ethnicity? Is this an organizational effect that is accounted for by the interaction of treatment with school-proportion minority? Also, we can discuss the exploratory results with experts, and consult earlier studies to establish a warrant for the effect being real (or not). Deeper exploration sharpens our understanding of the boundary conditions for observing a differential impact.
- More generally, we call for an approach to conducting randomized trials that, from the start, has built-in tests of contingencies for effects. This research agenda includes stating policy- theory- and practice-guided conjectures about conditions for a program effect being moderated, and designs powered to test these conjectures.
- Final note: We are not calling for elimination of replication, nor are we calling for an 'anything goes' approach. On the contrary, we are calling for even more rigor by advancing research agendas that assess the boundary conditions for observing effects.